Math 217: Topics in Applied Math
Optimal Transport
Spring 2022,
Instructor: Bo Li
Lectures: 1:00 – 1:50, MWF, AP&M 5402

Lecture 1. Monday, 3/28/2022

- Brief description of the course.
- Discrete optimal transport (OT)
  Monge's formulation

About optimal transport A subject of optimization, PDEs, the calculus of variations, probability, geometry, etc. with applications in economics, imaging science, molecular modeling, machine learning, etc.

About this course

- Applied math / comput. math aspects of optimal transport (OT)
- Introduction of the subject.
- Highlight some research areas / projects

Topics to be covered

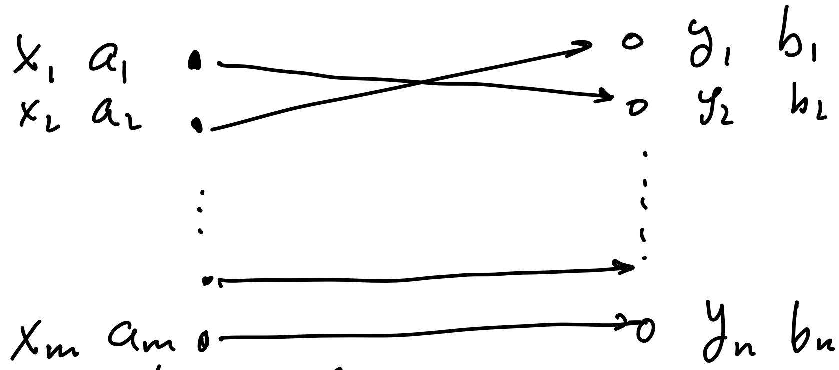
- Monge and Kantorovich formulations equivalence, duality, existence, etc. Wasserstein metric
- Discrete OT. regularization, numerical methods, stability, and convergence
- Gradient flow with Wasserstein metric. Fokker-Planck equation and other evolution equations
- Applications in machine learning and molecular dynamics

References See the course web page.

Expect Participation in class discussions. Possibly read some papers / presentations.

Part I Discrete Optimal Transport (OT)

Monge's formulation of discrete OT problem



Given/consider:

- $X = \{x_1, \dots, x_m\}$,

  $a_1, \dots, a_m \in (0,1)$, $\sum_i a_i = 1$

  $Y = \{y_1, \dots, y_n\}$

  $b_1, \dots, b_n \in (0,1)$, $\sum_j b_j = 1$.

  amount a_i of raw material at wherehouse x_i

  amount b_j of raw material to be transported to factory y_j

- Possible/feasible transport map $T: X \to Y$ such that

$$\boxed{b_j = \sum_{i:\, Tx_i = y_j} a_i \quad (j = 1, \dots, n)} \quad (*)$$

Denote

$$\mathcal{T} = \{ \text{all } T: X \to Y \text{ satisfying } (*) \}$$

- Cost function
  $c : X \times Y \longrightarrow [0, \infty)$
  $c(x_i, y_j)$: the cost to transport a unit product from x_i to y_j.
  e.g., $c(x_i, y_j) = \rho(x_i, y_j)$ if X, Y are subsets of a metric space with metric ρ.

The total cost for $T \in \mathcal{T}$ is
$$E[T] \triangleq \sum_{i=1}^{m} a_i\, c(x_i, T(x_i)).$$

<u>Monge's</u> <u>(discrete) OT problem</u>

Find $\hat{T} \in \mathcal{T}$ s.t.
$$E[\hat{T}] = \min_{T \in \mathcal{T}} E[T].$$
Call $\hat{T}$ an <u>optimal transport map</u>.

<u>Remarks</u> The constraint ($*$) is crucial.

- $\exists$ sol'n $\iff \mathcal{T} \neq \phi$.
  Sol'n may not be unique.

<u>Exercise</u> Example of nonuniqueness?

- Since each $b_j > 0$, we have T is onto (i.e., surjective). Hence, $m \geq n$.

Consider the case $n = m$.

Each $T \in \mathcal{T}$ is a bijection (1-1, onto)

If all a_i are distinct, then T is unique. $Tx_i = y_j$ with $b_j = a_i$.

If some a_i are the same, then the problem can be decomposed into some subproblems:

$$
\begin{array}{llll}
n_1 \left\{ \begin{matrix} a_1 & x_1' \\ \vdots & \vdots \\ a_1 & x_{n_1}' \end{matrix} \right\} & \xrightarrow{T_1'} & \left\{ \begin{matrix} a_1 & y_1' \\ \vdots & \vdots \\ a_1 & y_{n_1}' \end{matrix} \right\} \\[2ex]
n_2 \left\{ \begin{matrix} a_2 & x_{n_1+1}' \\ \vdots & \vdots \\ a_2 & x_{n_2} \end{matrix} \right\} & \xrightarrow{T_2'} & \left\{ \begin{matrix} a_2 & y_{n_1+1}' \\ \vdots & \vdots \\ a_2 & y_{n_2}' \end{matrix} \right\} \\[1ex]
\quad \vdots \qquad \vdots & & \\[1ex]
n_k \left\{ \begin{matrix} a_k & x_{n_{k-1}+1}' \\ \vdots & \vdots \\ a_k & x_n' \end{matrix} \right\} & \xrightarrow{T_k'} & \left\{ \begin{matrix} a_k & y_{n_{k-1}+1}' \\ \vdots & \vdots \\ a_k & y_n' \end{matrix} \right\}
\end{array}
$$

$$\sum_{l=1}^{k} n_l = n. \qquad \text{Find } T_j'.$$

Two steps:
- ⊙ Relabel x_i and y_j
- ⊙ For each group of x_i with same a-value, solve the OT problem

The optimal assignment problem (as a Monge's prob of discret OT)

Given: $X = Y = \{1, 2, \cdots, n\}$.

$a_i = b_i = \frac{1}{n}$ $(i = 1, \cdots, n)$.

$C = [C_{ij}] \in \mathbb{R}^{n \times n}$, all $C_{ij} \geq 0$.

Notation: $\mathbb{R}^{m \times n} = \{\text{all } m \times n \text{ real matrices}\}$.

$S_n = \{\text{all permutations of } (1\ 2 \cdots n)\}$

e.g. $S_3 = \{(123), (132), (213), (231), (312), (321)\}$.

A permutation is a bijection.

$$\boxed{\begin{array}{l} \text{The (optimal) assignment prob. (in Monge's form)} \\ \text{Find } \hat{\sigma} \in S_n \text{ s.t.} \\ \qquad \hat{\sigma} = \arg \min\limits_{\sigma \in S_n} \frac{1}{n} \sum\limits_{i=1}^{n} C_{i,\sigma(i)} \end{array}}$$

- Hence the cost function is
$$c(i,j) = C_{ij}, \quad i,j = 1, \cdots, n.$$
- $c(x_i, T(x_i)) = c(x_i, y_{\sigma_i}) = C_{i,\sigma_i}$.
- Sol'ns exist but often nonunique. But $\mathrm{Card}(S_n) = n!$

$5! = 120$, $8! = 40{,}320$, $10! = 3{,}628{,}800$

$12! = 479{,}001{,}600$, $25! = 1{,}551 \times 10^{25}$, $70! = 1{,}198 \times 10^{100}$

We will revisit the problem later.

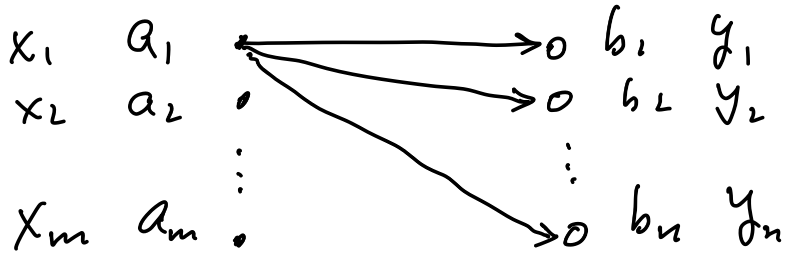
The case $m > n$. More complicated.

But, interesting?

<u>Exercise</u> Solve this problem.

Kantorovich's formulation of the (discrete) OT problem

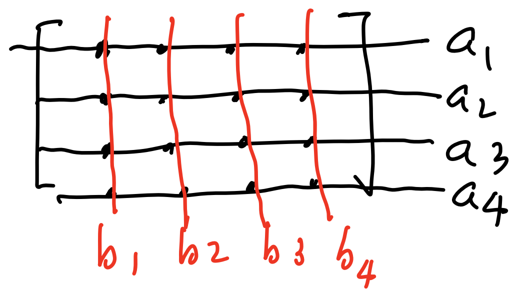
- Allowing the split of a_i into pieces
- Probabilistic approach



Given/consider:

- $X = \{x_1, \cdots, x_m\}$, $Y = \{y_1, \cdots, y_n\}$
  $a_i \geq 0$, $\sum_i a_i = 1$, $b_j \geq 0$, $\sum_j b_j = 1$
- $c_{ij}\ (\geq 0)$ = cost for transporting a unit product from x_i to y_j
  $C = [c_{ij}] \in \mathbb{R}^{m \times n}$: cost matrix
- feasible transport plans
  $P = [p_{ij}] \in \mathbb{R}^{m \times n}$, all $p_{ij} \geq 0$.
  p_{ij} = amount of product i (i.e., part of a_i) at x_i transported to y_j, becoming part of b_j.

$$\boxed{\begin{aligned}&\sum_{i=1}^{m} p_{ij} = b_j \quad (j=1,\cdots,n), \ \text{col. sum of } P = b\\ &\sum_{j=1}^{n} p_{ij} = a_i \quad (i=1,\cdots,m), \ \text{Row sum of } P = a\end{aligned}} \quad (*)$$



Note: $\sum_{i,j} p_{ij} = 1$.

<u>The feasible set of transport plans:</u>

$$\mathcal{A}(a,b) \triangleq \Big\{ P = [p_{ij}] \in \mathbb{R}^{m\times n} : \text{all } p_{ij} \geq 0, \ \sum_i p_{ij} = b_j, \forall j, \ \sum_j p_{ij} = a_i, \forall i \Big\}.$$

Call $E[P] := \sum_{i,j} p_{ij} c_{ij}$ the total cost of the plan P.

<u>Kantorovich's formulation</u>

$$\boxed{\begin{aligned}&\text{Find } \hat{P} \in \mathcal{A}(a,b) \text{ s.t.}\\ &\qquad \hat{P} = \arg\min_{P \in \mathcal{A}(a,b)} \sum_{i,j} p_{ij} c_{ij}.\end{aligned}}$$